

Figure 1. *H. grisea* GSHE nucleotide sequence with *putative* introns bold & underlined.

ATGCATACCTTCTCCAAGCTCCTCGTCCTGGGCTCTGCCGTCCAGTCTGCCCTCGGGCGGCCTCACGGCTCT  
 TCGCGTCTCCAGGAACGCGCTGCCGTTGATACCTTCATCAACACCGAGAAGCCCATCGCATGGAACAAGCTG  
 CTCGCCAACATCGGCCCTAACGGCAAAGCCGCTCCCGGTGCCGCCGCCGGCGTTGTGATTGCCAGCCCTTCC  
 AGGACGGACCCCTCCTT**GTACGTGGTGGCATGGAATGGACCCAAGAGACTGGTTTTAGATGAAAGAGAGTTTC**  
**TGCTAACCGCCACACCCAGACTTCTTCACCTGGACCCGCGATGCCGCCCTGGTCCTCACCGGCATCATCGAG**  
 TCCCTTGGCCACAACACTACAACACCACCCTGCAGACCGTCATCCAGAACTACGTGCGCTCGCAGGCCAAGCTG  
 CAGCAGGTCTCGAACCCCTCGGGAACCTTCGCCGACGGCTCGGGTCTCGGTGAGGCCAAGTTCAATGTCGAC  
 CTCACTGCCTTCACTGGCGAATGGGGTCGCCCTCAGAGGGACGGCCCCGCCCTGCGCGCCATCGCTCTCATC  
 CAGTACGCCAAGTGGCTGATCGCCAACGGCTACAAGAGCACGGCCAAGAGCGTCGTCTGGCCCGTCGTCAAG  
 AACGATCTCGCCTACACGGCCCAGTACTGGAACGAGACCGGCTTCGATCTCTGGGAGGAGGTCCCCGGCAGC  
 TCGTTCTTTACCATCGCCAGCTCTCACAGGG**GTGAGTCATTTATTGTTTCAGTGTTTTCTCATTTGAATAATTA**  
**CCGGAATGCCACTGACGCCAAACAGCTCTGACTGAGGGTGCTTACCTCGCCGCTCAGCTCGACACCGAGTGC**  
 CGCGCCTGCACGACCGTCGCCCTCAGGTTCTGTGCTTCCAGCAGGCCTTCTGGAACCTCAAGGGCAACTAT  
 GTCGTCTCCAACAG**GTAAGATCCCTACACCAACAAAAAATCGAAAAGGAACGTTAGCTGACCCTTCTAGTC**  
 AACGGCGGCGAGTATCGCTCCGGCAAGGACGCCAACTCGATCCTGGCGTCCATCCACAACCTTCGACCCTGAG  
 GCCGGCTGCGACAACCTGACCTTCCAGCCCTGCAGCGAGCGCGCCCTGGCCAACCACAAGGCCTATGTCGAC  
 TCGTTCCGCAACCTCTAGCCATCAACAAGGGCATCGCCAGGGCAAGGCCGTTGCCGTGCGCCGCTACTCG  
 GAGGATGTCTACTACAACGGCAACCCGTGGTACCTGGCCAACCTTGGCCGCCGCGAGCAGCTCTACGACGCC  
 ATCTACGTGTGGAACAAGCAGGGCTCCATCACCGTGACCTCGGTCTCCCTGCCCTTCTTCCGCGACCTTGTC  
 TCGTCGGTCAGCACCGGCACCTACTCCAAGAGCAGCTCGACCTTACCAACATCGTCAACGCCGTCAAGGCC  
 TACGCCGACGGCTTCATCGAGGTGGCGGCCAAGTACACCCCGTCCAACGGCGCGCTCGCCGAGCAGTACGAC  
 CGCAACACGGGCAAGCCCGACTCGGCCGCCGACCTGACGTGGTCTGACTCGGCCTTCTCTCGGCCATCGAC  
 CGCCGCGCGGGTCTCGTCCCCCGAGCTGGCGGGCCAGCGTGGCCAAGAGCCAGCTGCCGTCCACCTGCTCG  
 CGCATCGAGGTGCGCCGGCACCTACGTGCGGCCACGAGCACCTCGTTCCCGTCCAAGCAGACCCCGAACCC  
 TCCGCGGCGCCCTCCCCGTCCCCCTACCCGACCGCCTGCGCGGACGCTAGCGAGGTGTACGTCACCTTCAAC  
 GAGCGCGTGTGACCGCGTGGGGCGAGACCATCAAGGTGGTGGGCAACGTGCCGGCGCTGGGGAACTGGGAC  
 ACGTCCAAGGCGGTGACCCTGTGCGCCAGCGGGTACAAGTCGAATGATCCCCTCTGGAGCATCACGGTGCCC  
 ATCAAGGCGACGGGCTCGGCCGTGAGTACAAGTATATCAAGGTGGCACCAACGGGAAGATTACTTGGGAG  
 TCGGACCCCAACAGGAGCATTACCCTGCAGACGGCGTCGTCTGCGGGCAAGTGCGCCGCGCAGACGGTGAAT  
 GATTCGTGGCGTTAA

Figure 2A. *H. grisea* GSHE protein sequence with *putative* signal sequence underlined.

```

M H T F S K L L V L G S A V Q S A L G R P H G S S R L Q E R A A V D T F I N T E
K P I A W N K L L A N I G P N G K A A P G A A A G V V I A S P S R T D P P Y F F
T W T R D A A L V L T G I I E S L G H N Y N T T L Q T V I Q N Y V A S Q A K L Q
Q V S N P S G T F A D G S G L G E A K F N V D L T A F T G E W G R P Q R D G P P
L R A I A L I Q Y A K W L I A N G Y K S T A K S V V W P V V K N D L A Y T A Q Y
W N E T G F D L W E E V P G S S F F T I A S S H R A L T E G A Y L A A Q L D T E
C R A C T T V A P Q V L C F Q Q A F W N S K G N Y V V S N I N G G E Y R S G K D
A N S I L A S I H N F D P E A G C D N L T F Q P C S E R A L A N H K A Y V D S F
R N L Y A I N K G I A Q G K A V A V G R Y S E D V Y Y N G N P W Y L A N F A A A
E Q L Y D A I Y V W N K Q G S I T V T S V S L P F F R D L V S S V S T G T Y S K
S S S T F T N I V N A V K A Y A D G F I E V A A K Y T P S N G A L A E Q Y D R N
T G K P D S A A D L T W S Y S A F L S A I D R R A G L V P P S W R A S V A K S Q
L P S T C S R I E V A G T Y V A A T S T S F P S K Q T P N P S A A P S P S P Y P
T A C A D A S E V Y V T F N E R V S T A W G E T I K V V G N V P A L G N W D T S
K A V T L S A S G Y K S N D P L W S I T V P I K A T G S A V Q Y K Y I K V G T N
G K I T W E S D P N R S I T L Q T A S S A G K C A A Q T V N D S W R

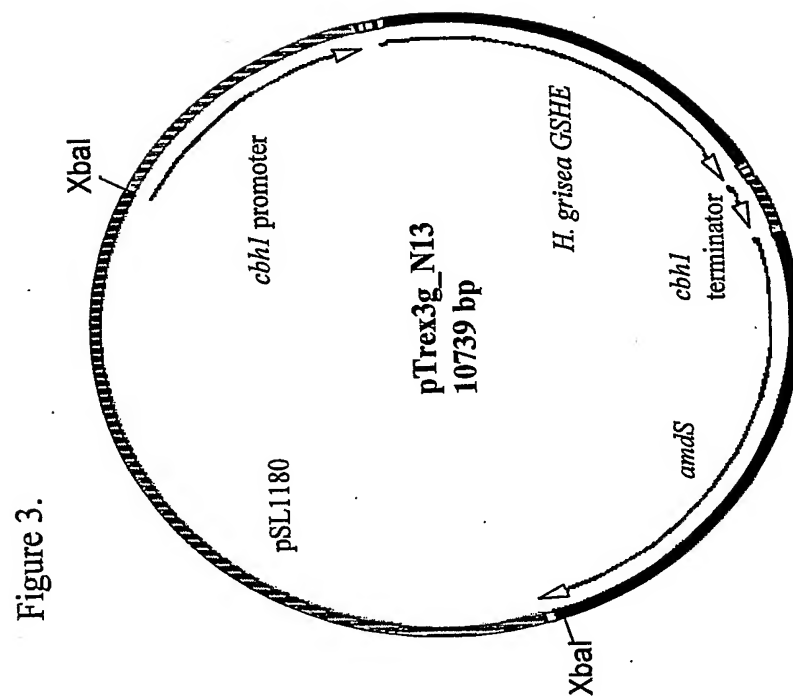
```

Figure 2B. *H. grisea* Mature GSHE protein sequence

```

A A V D T F I N T E K P I A W N K L L A N I G P N G K A A P G A A A G V V I A S
P S R T D P P Y F F T W T R D A A L V L T G I I E S L G H N Y N T T L Q T V I Q
N Y V A S Q A K L Q Q V S N P S G T F A D G S G L G E A K F N V D L T A F T G E
W G R P Q R D G P P L R A I A L I Q Y A K W L I A N G Y K S T A K S V V W P V V
K N D L A Y T A Q Y W N E T G F D L W E E V P G S S F F T I A S S H R A L T E G
A Y L A A Q L D T E C R A C T T V A P Q V L C F Q Q A F W N S K G N Y V V S N I
N G G E Y R S G K D A N S I L A S I H N F D P E A G C D N L T F Q P C S E R A L
A N H K A Y V D S F R N L Y A I N K G I A Q G K A V A V G R Y S E D V Y Y N G N
P W Y L A N F A A A E Q L Y D A I Y V W N K Q G S I T V T S V S L P F F R D L V
S S V S T G T Y S K S S S T F T N I V N A V K A Y A D G F I E V A A K Y T P S N
G A L A E Q Y D R N T G K P D S A A D L T W S Y S A F L S A I D R R A G L V P P
S W R A S V A K S Q L P S T C S R I E V A G T Y V A A T S T S F P S K Q T P N P
S A A P S P S P Y P T A C A D A S E V Y V T F N E R V S T A W G E T I K V V G N
V P A L G N W D T S K A V T L S A S G Y K S N D P L W S I T V P I K A T G S A V
Q Y K Y I K V G T N G K I T W E S D P N R S I T L Q T A S S A G K C A A Q T V N
D S W R

```



## FIGURE 4A

AAGCTTACTAGTACTTCTCGAGCTCTGTACATGTCCGGTCGCGACGTACGCGTATCGATGGCGCCAGCTG  
CAGGCGGCCCGCCTGCAGCCACTTGCAGTCCCGTGGAATTCTCACGGTGAATGTAGGCCTTTTGTAGGGTA  
GGAATTGTCACTCAAGCACCCCCAACCTCCATTACGCCTCCCCCATAGAGTTCCCAATCAGTGAGTCATG  
GCACTGTTTCTCAAATAGATTGGGGAGAAGTTGACTTCCGCCCAGAGCTGAAGGTCGCACAACCGCATGAT  
ATAGGGTCGGCAACGGCAAAAAGCACGTGGCTCACCGAAAAGCAAGATGTTTGCGATCTAACATCCAGG  
AACCTGGATACATCCATCATCACGCACGACCACTTTGATCTGCTGGTAAACTCGTATTCGCCCTAAACCG  
AAGTGCGTGGTAAATCTACACGTGGGCCCTTTCGGTATACTGCGTGTGTCTTCTCTAGGTGCCATTCTT  
TTCCCTTCTCTAGTGTTGAATTGTTTGTGTTGGAGTCCGAGCTGTAACCTACCTCTGAATCTCTGGAGAA  
TGGTGGACTAACGACTACCGTGCACCTGCATCATGTATATAATAGTGATCCTGAGAAGGGGGGTTTGGAG  
CAATGTGGGACTTTGATGGTCATCAAACAAAGAACGAAGACGCCTCTTTTGCAAAGTTTTGTTTCGGCTA  
CGGTGAAGAACTGGATACTTGTTGTGTCTTCTGTGTATTTTTGTGGCAACAAGAGGCCAGAGACAATCTA  
TTCAAACACCAAGCTTGCTCTTTTGAGCTACAAGAACCTGTGGGGTATATATCTAGAGTTGTGAAGTCGG  
TAATCCCGCTGTATAGTAATACGAGTCGCATCTAAATACTCCGAAGCTGCTGCGAACCCGGAGAATCGAG  
ATGTGCTGGAAAGCTTCTAGCGAGCGGCTAAATTAGCATGAAAGGCTATGAGAAATTCTGGAGACGGCTT  
GTTGAATCATGGCGTTCCATTCTTCGACAAGCAAAGCGTTCCGTGCGAGTAGCAGGCACTCATTTCCCGAA  
AAAACCTCGGAGATTCCCTAAGTAGCGATGGAACCGGAATAATATAATAGGCAATACATTGAGTTGCCTCGA  
CGGTTGCAATGCAGGGGTACTGAGCTTGACATAACTGTTCCGTACCCACCTCTTCTCAACCTTTGGCG  
TTTCCCTGATTACGCGTACCCGTACAAGTCGTAATCACTATTAACCCAGACTGACCGGACGTGTTTTGCC  
CTTCATTTGGAGAAATAATGTCATTGCGATGTGTAATTTGCCTGCTTGACCGACTGGGGCTGTTTCAAGC  
CCGAATGTAGGATTGTTATCCGAACCTCTGCTCGTAGAGGCATGTTGTGAATCTGTGTGCGGCAGGACACG  
CCTCGAAGGTTACGGCAAGGGAAACCCGATAGCAGTGTCTAGTAGCAACCTGTAAAGCCGCAATGCA  
GCATCACTGGAAAATACAAACCAATGGCTAAAAGTACATAAGTTAATGCCTAAAGAAGTCATATACCAGC  
GGCTAATAATTGTACAATCAAGTGGCTAAACGTACCGTAATTTGCCAACGGCTTGTGGGGTTGCAGAAGC  
AACGGCAAAGCCCCACTTCCCCAGTTTGTCTTCACTCAGTCCAATCTCAGCTGGTGATCCCCCAATT  
GGGTCGCTTGTTTGTTCGGTGAAGTGAAAGAAGACAGAGGTAAGAATGTCTGACTCGGAGCGTTTTGCA  
TACAACCAAGGGCAGTGATGGAAGACAGTGAAATGTTGACATTCAAGGAGTATTTAGCCAGGGATGCTTG  
AGTGTATCGTGTAAGGAGGTTTGTCTGCCGATACGACGAATACTGTATAGTCACTTCTGATGAAGTGGTC  
CATATTGAAATGTAAAGTCGGCACTGAACAGGCAAAAGATTGAGTTGAAACTGCCTAAGATCTCGGGCCC  
TCGGGCCTTCGGCCTTTGGGTGTACATGTTTGTGCTCCGGGCAAATGCAAAGTGTTGGTAGGATCGAACAC  
ACTGCTGCCTTTACCAAGCAGCTGAGGGTATGTGATAGGCAAATGTTTCAGGGGGCCACTGCATGGTTTCGA  
ATAGAAAGAGAAGCTTAGCCAAGAACAATAGCCGATAAAGATAGCCTCATTAACGGAATGAGCTAGTAG

## FIGURE 4B

GCAAAGTCAGCGAATGTGTATATATAAAGGTTTCGAGGTCCGTGCCTCCCTCATGCTCTCCCCATCTACTC  
ATCAACTCAGATCCTCCAGGAGACTTGTACACCATCTTTTGAGGCACAGAAACCCAATAGTCAACCATCA  
CAAGTTTGTACAAAAAAGCAGGCTCCGCGGCCGCCCTTCAACATGCATACCTTCTCCAAGCTCCTCGT  
CCTGGGCTCTGCCGTCCAGTCTGCCCTCGGGCGGCCCTCACGGCTCTTCGCGTCTCCAGGAACGCGCTGCC  
GTTGATACCTTCATCAACACCGAGAAGCCCATCGCATGGAACAAGCTGCTCGCCAACATCGGCCCTAACG  
GCAAAGCCGCTCCCGGTGCCGCCGCCGGCGTTGTGATTGCCAGCCCTTCCAGGACGGACCCTCCTTGTAC  
GTGGTGGCATGGAATGGACCCAAGAGACTGGTTTTAGATGAAAGAGAGTTTCTGCTAACCGCCACACCCA  
GACTTCTTCACCTGGACCCGCGATGCCGCCCTGGTCCTCACCGGCATCATCGAGTCCCTTGGCCACAAC  
ACAACACCACCCTGCAGACCGTCATCCAGAACTACGTCGCGTCGCAGGCCAAGCTGCAGCAGGTCTCGAA  
CCCCTCGGGAACCTTCGCCGACGGCTCGGGTCTCGGTGAGGCCAAGTTCAATGTCGACCTCACTGCCTTC  
ACTGGCGAATGGGGTCGCCCTCAGAGGGACGGCCCGCCCCTGCGCGCCATCGCTCTCATCCAGTACGCCA  
AGTGGCTGATCGCCAACGGCTACAAGAGCACGGCCAAGAGCGTCGTCTGGCCCGTCGTCAAGAACGATCT  
CGCCTACACGGCCCAGTACTGGAACGAGACCGGCTTCGATCTCTGGGAGGAGGTCCCCGGCAGCTCGTTC  
TTTACCATCGCCAGCTCTCACAGGGGTGAGTCATTTATTGTTTCAGTGTTTTCTCATTTGAATAATTACCGG  
AATGCCACTGACGCCAAACAGCTCTGACTGAGGGTGCTTACCTCGCCGCTCAGCTCGACACCGAGTGCCG  
CGCCTGCACGACCGTCGCCCCCTCAGGTTCTGTGCTTCCAGCAGGCCTTCTGGAACCTCAAGGGCAACTAT  
GTCGTCTCCAACAGTAAGATCCCTACACCAACAAAAAAATCGAAAAGGAACGTTAGCTGACCCCTTCTAG  
TCAACGGCGGCGAGTATCGCTCCGGCAAGGACGCCAACTCGATCCTGGCGTCCATCCACAACCTTCGACCC  
TGAGGCCCGGCTGCGACAACCTGACCTTCCAGCCCTGCAGCGAGCGCGCCCTGGCCAACCACAAGGCCTAT  
GTCGACTCGTTCCGCAACCTCTACGCCATCAACAAGGGCATCGCCCAGGGCAAGGCCGTTGCCGTGCGCC  
GCTACTCGGAGGATGTCTACTACAACGGCAACCCGTGGTACCTGGCCAACCTTTGCCGCCGCCGAGCAGCT  
CTACGACGCCATCTACGTGTGGAACAAGCAGGGCTCCATCACCGTGACCTCGGTCTCCCTGCCCTTCTTC  
CGCGACCTTGTCTCGTTCGGTCAGCACCGGCACCTACTCCAAGAGCAGCTCGACCTTACCAACATCGTCA  
ACGCCGTCAAGGCCTACGCCGACGGCTTCATCGAGGTGGCGGCCAAGTACACCCCGTCCAACGGCGCGCT  
CGCCGAGCAGTACGACCGCAACACGGGCAAGCCCGACTCGGCCGCCGACCTGACGTGGTTCGTACTCGGCC  
TTCCTCTCGGCCATCGACCGCCGCGCGGGTCTCGTCCCCCGAGCTGGCGGGCCAGCGTGGCCAAGAGCC  
AGCTGCCGTCCACCTGCTCGCGCATCGAGGTGCGCGGCACCTACGTGCGCGCCACGAGCACCTCGTTCCC  
GTCCAAGCAGACCCCGAACCCTCCGCGGCGCCCTCCCCGTCCCCCTACCCGACCGCCTGCGCGGACGCT  
AGCGAGGTGTACGTACCTTCAACGAGCGCGTGTGACCGCGTGGGGCGAGACCATCAAGGTGGTGGGCA  
ACGTGCCGGCGCTGGGGAACTGGGACACGTCCAAGGCGGTGACCCTGTGCGCCAGCGGGTACAAGTCGAA  
TGATCCCCCTCTGGAGCATCACGGTGCCCATCAAGGCGACGGGCTCGGCCGTGCAGTACAAGTATATCAAG

## FIGURE 4C

GTCGGCACCAACGGGAAGATTACTTGGGAGTCGGACCCCAACAGGAGCATTACCCTGCAGACGGCGTCGT  
CTGCGGGCAAGTGCGCCGCGCAGACGGTGAATGATTCGTGGCGTTAAAAGGGTGGGCGCGCCGACCCAGC  
TTTCTTGTAACAAAGTGGTGATCGCGCCAGCTCCGTGCGAAAGCCTGACGCACCGGTAGATTCTTGGTGAG  
CCCGTATCATGACGGCGGGGAGCTACATGGCCCCGGGTGATTTATTTTTTTTGTATCTACTTCTGACC  
CTTTTCAAATATACGGTCAACTCATCTTTCCTGAGATGCGGCCTGCTTGGTATTGCGATGTTGTCAGC  
TTGGCAAATTGTGGCTTTTCGAAAACACAAAACGATTCCTTAGTAGCCATGCATTTTAAGATAACGGAATA  
GAAGAAAGAGGAAATTAATAAAAAAAAAAAAAAAAAACAAACATCCCGTTCATAACCCGTAGAATCGCCGCTCTT  
CGTGTATCCAGTACCAGTTTATTTTGAATAGCTCGCCCGCTGGAGAGCATCCTGAATGCAAGTAACAAC  
CGTAGAGGCTGACACGGCAGGTGTTGCTAGGGAGCGTCGTGTTCTACAAGGCCAGACGTCTTCGCGGTTG  
ATATATATGTATGTTTGACTGCAGGCTGCTCAGCGACGACAGTCAAGTTCGCCCTCGCTGCTTGTGCAAT  
AATCGCAGTGGGGAAGCCACACCGTGACTCCCATCTTTCAGTAAAGCTCTGTTGGTGTTCATCAGCAATA  
CACGTAATTTAACTCGTTAGCATGGGGCTGATAGCTTAATTACCGTTTACCAGTGCCATGGTTCTGCAG  
CTTTCCTTGCCCCGTAAAATTCGGCGAAGCCAGCCAATCACCAGCTAGGCACCAGCTAAACCCTATAATT  
AGTCTCTTATCAACACCATCCGCTCCCCCGGGATCAATGAGGAGAATGAGGGGGATGCGGGGGCTAAAGAA  
GCCTACATAACCTCATGCCAACTCCCAGTTTACACTCGTCGAGCCAACATCCTGACTATAAGCTAACAC  
AGAATGCCTCAATCCTGGGAAGAACTGGCCGCTGATAAGCGCGCCCGCCTCGCAAAAACCATCCCTGATG  
AATGGAAAGTCCAGACGCTGCCTGCGGAAGACAGCGTTATTGATTTCCCAAAGAAATCGGGGATCCTTTC  
AGAGGCCGAAGTGAAGATCACAGAGGCCTCCGCTGCAGATCTTGTGTCCAAGCTGGCGGGCCGGAGAGTTG  
ACCTCGGTGGAAGTTACGCTAGCATTCGTGTAACGGGCAGCAATCGCCCAGCAGTTAGTAGGGTCCCCTC  
TACCTCTCAGGGAGATGTAACAACGCCACCTTATGGGACTATCAAGCTGACGCTGGCTTCTGTGCAGACA  
AACTGCGCCACAGAGTTCTTCCCTGACGCCGCTCTCGCGCAGGCAAGGGAACCTCGATGAATACTACGCAA  
AGCACAAGAGACCCGTTGGTCCACTCCATGGCCTCCCCATCTCTCTCAAAGACCAGCTTCGAGTCAAGGT  
ACACCGTTGCCCTAAGTCGTTAGATGTCCCTTTTTGTGAGCTAACATATGCCACCAGGGCTACGAAACA  
TCAATGGGCTACATCTCATGGCTAAACAAGTACGACGAAGGGGACTCGGTTCTGACAACCATGCTCCGCA  
AAGCCGGTGCCGTCTTCTACGTCAAGACCTCTGTCCCGCAGACCCTGATGGTCTGCGAGACAGTCAACAA  
CATCATCGGGCGCACCGTCAACCCACGCAACAAGAACTGGTCGTGCGGCGGCAGTTCTGGTGGTGAGGGT  
GCGATCGTTGGGATTTCGcRVTGGTGGCGTCATCGGTGTAGGAACGGATATCGGTGGCTCGATTGAGTGC  
CGGCCGCGTTCAACTTCCTGTACGGTCTAAGGCCGAGTCATGGGCGGCTGCCGTATGCAAAGATGGCGAA  
CAGCATGGAGGGTCAGGAGACGGTGCACAGCGTTGTGGGCCGATTACGCACTCTGTTGAGGGTGAGTCC  
TTCGCCTCTTCCTTCTTTTCCTGCTCTATACCAGGCCTCCACTGTCCTCCTTTCTTGCTTTTTATACTAT  
ATACGAGACCGGCAGTCACTGATGAAGTATGTTAGACCTCCGCCTCTTCACCAAATCCGTCCCTCGGTGAG

## FIGURE 4D

GAGCCATGGAAATACGACTCCAAGGTCATCCCCATGCCCTGGCGCCAGTCCGAGTCGGACATTATTGCCT  
CCAAGATCAAGAACGGCGGGCTCAATATCGGCTACTACAACTTCGACGGCAATGTCTTCCACACCCTCC  
TATCCTGCGCGGCGTGGAACACCGTCGCCGCACTCGCCAAAGCCGGTCACACCGTGACCCCGTGGACG  
CCATACAAGCACGATTTTCGGCCACGATCTCATCTCCCATATCTACGCGGCTGACGGCAGC<sub>6</sub>RVGCCGACG  
TAATGCGCGATATCAGTGCATCCGGCGAGCCGGCGATTCCAAATATCAAAGACCTACTGAACCCGAACAT  
CAAAGCTGTTAACATGAACGAGCTCTGGGACACGCATCTCCAGAAGTGGAATTACCAGATGGAGTACCTT  
GAGAAATGGCGGGAGGCTGAAGAAAAGGCCGGGAAGGAACTGGACGCCATCATCGCGCCGATTACGCCTA  
CCGCTGCGGTACGGCATGACCAGTTCCGGTACTATGGGTATGCCTCTGTGATCAACCTGCTGGATTTAC  
GAGCGTGTTGTTCCGGTTACCTTTGCGGATAAGAACATCGATAAGAAGAATGAGAGTTTCAAGGCGGTT  
AGTGAGCTTGATGCCCTCGTGCAGGAAGAGTATGATCCGGAGGCGTACCATGGGGCACC GGTTGCAGTGC  
AGGTTATCGGACGGAGACTCAGTGAAGAGAGGACGTTGGCGATTGCAGAGGAAGTGGGGAAGTTGCTGGG  
AAATGTGGTGACTCCATAGCTAATAAGTGTGAGATAGCAATTTGCACAAGAAATCAATACCAGCAACTGT  
AAATAAGCGCTGAAGTGACCATGCCATGCTACGAAAGAGCAGAAAAAACCTGCCGTAGAACCGAAGAGA  
TATGACACGCTTCCATCTCTCAAAGGAAGAATCCCTTCAGGGTTGCGTTTCCAGTCTAGACACGTATAAC  
GGCACAAGTGTCTCTCACCAAATGGGTTATATCTCAAATGTGATCTAAGGATGGAAAGCCCAGAATATCG  
ATCGCGCGCAGATCCATATATAGGGCCCGGGTTATAATTACCTCAGGTCGACGTCCCATGGCCATTTCGAA  
TTCGTAATCATGGTCATAGCTGTTTCCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGA  
GCCGGAAGCATAAAGTGTAAGCCTGGGGTGCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCT  
CACTGCCCCGCTTTCCAGTCGGGAAACCTGTGCTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAG  
AGGCGGTTTTCGCTATTGGGCGCTCTTCCGCTTCCCTCGCTCACTGACTCGCTGCGCTCGGTTCGTTCCGGCTG  
CGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAA  
AGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA  
TAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGA  
CTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCCTCTCCTGTTCCGACCCTGCCGCTTA  
CCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCT  
CAGTTCGGTGTAGGTTCGTTCCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTACGCCCGACCGCTGC  
GCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCA  
CTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCCTGAAGTGGTGGCCTAACTA  
CGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTT  
GGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTA  
CGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGA

## FIGURE 4E

AAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAA  
AAATGAAGTTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCA  
GTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT  
AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCG  
GCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTAT  
CCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTTCGCCAGTTAATAGTTTGCG  
CAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTACAGCTCC  
GGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTC  
CTCCGATCGTTGTCAGAAGTAAGTTGGCCGCGAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTC  
TCTTACTGTGTCGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAA  
TAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAA  
CTTTAAAAGTGCTCATCATTTGGAACGTTCTTCGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAG  
ATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCT  
GGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATAC  
TCATACTCTTCCTTTTTCAATATTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGATACATATT  
TGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC  
TAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGC  
GTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGC  
GGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAAC  
TATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATAAAAATTGTAAACGTTAATATTTTGTAAAT  
TCGCGTTAAATTTTTGTAAATCAGCTCATTTTTTAAACCAATAGGCCGAAATCGGCAAAATCCCTTATAA  
ATCAAAAGAATAGCCCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAAC  
GTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCA  
AATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCTTAAAGGGAGCCCCGATTTAG  
AGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGAAGAAAGCGAAAGGAGCGGGCGCTAGG  
GCGCTGGCAAGTGTAGCGGTACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGG  
GCGCGTACTATGGTTGCTTTGACGTATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAAATACCGC  
ATCAGGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTAT  
TACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAAGTTGGGTAACGCCAGGTTTTCCAGTC  
ACGACGTTGTAAAACGACGGCCAGTGCCC



FIGURE 5: Genomic Nucleotide Sequence of *Aspergillus kawachi* GSHE

ATGTCGTTCCGATCTCTTCTCGCCCTGAGCGGCCTTGTCTGCTCGGGGTTGGCAAGTGTGATTTC  
CAAGCGCGCGACCTTGGATTTCGTGGTTGAGCAACGAAGCGACCGTGGCCCGTACTGCGATCCTGA  
ATAACATCGGGGCGGACGGTGCTTGGGTGTGCGGCGCGGACTCTGGCATTGTCTGTTGCCAGTCCC  
AGCACCGATAACCCGGACTGTATGTTTTGAGTTCGGATTATGAATGTGTCTTGGTTGATTGATGC  
TGACTGGCGTGTCTTTTGATGATTGTAGACTTCTACACCTGGACTCGCGACTCTGGTCTCGTCAT  
CAAGACCCCTCGTCGACCTCTTCCGCAATGGAGATACTGATCTCCTTTCCACCATTGAGCACTACA  
TCTCCTCTCAGGCAATTATTCAGGGTGTGAGTAACCCCTCTGGTGATCTGTCCAGCGGTGGTCTT  
GGTGAGCCCAAGTTCAATGTCGATGAGACTGCCTACACCGGTTCTTGGGGACGGCCGAGCGTGA  
TGGTCCTGCCCTGAGAGCAACTGCTATGATCGGCTTTGGGCAGTGGCTGCTTGTATGTTCTCCAC  
CTCCTTGCGTCTGATCTGCAACATATGTAGCCGACTGGTCAGGACAATGGCTACACCAGCGCTGC  
AACAGAGATTGTTTGGCCCTCGTTAGGAACGACCTGTCTGATGTGGCTCAGTACTGGAACCAGA  
CGGGATATGGTGTGTTTGATTGATCGGGGTTCAAGGGTGTGTTGTGCATCGGAGCTAACTTCGCGG  
TCGCAGATCTCTGGGAAGAAGTTAATGGCTCGTCCTTCTTCACTATTGCCGTGCAACACCGCGCC  
CTCGTCGAAGGTAGTGCCTTCGCGACGGCCGTCGGCTCGTCCTGCTCCTGGTGTGATTTCGAGGC  
ACCTCAGATTCTCTGTTACTTGCAGTCCTTCTGGACCGGCAGCTACATCCTGGCCAACCTTTGACA  
GCAGCCGTTCCGGCAAGGACACAAAACACCCTCCTGGGAAGCATCCACACCTTTGATCCTGAGGCT  
GGATGCGACGACTCCACCTTCCAGCCCTGCTCCCCGCGTGCGCTCGCCAACCATAAGGAGGTTGT  
AGACTCTTTCCGCTCGATCTATACTCTCAACGATGGTCTCAGTGACAGTGAGGCGGTTGCGGTCG  
GTCGGTACCCTGAGGATAGCTACTACAACGGCAACCCGTGGTTCCTGTGCACCTTGGCTGCCGCG  
GAACAGCTGTACGATGCTCTGTACCAGTGGGACAAGCAGGGGTCGTTGGAGATCACAGACGTGTC  
ACTTGACTTCTTCAAGGCTCTGTACAGTGGTGTGCCACCGGCACGTACTCTTCGTCCAGCTCGA  
CCTATAGCAGCATTGTGAGTGCCGTCAAGACTTTCGCTGATGGTTTTGTCTTCTATTGTGGTAAGT  
CTACGCTAGACGAGCGCTCATATTTACAGAGGGTGCCTACTAACAGGATTAGGAACTCACGCCG  
CAAGCAACGGCTCTCTGTCTGAGCAATTCGACAAGTCTGATGGCGACGAGCTTTCTGCTCGCGAT  
CTGACCTGGTCTTACGCTGCTCTGCTGACCGCCAACAACCGTCGTAATTCTGTCTGCCCCCGTC  
TTGGGGTGAGACCTCTGCCAGCAGCGTGCCCGGCACCTGTGCGGCTACCTCTGCCTCTGGTACCT  
ACAGCAGTGTGACCGTCACCTCGTGGCCGAGCATCGTGGCTACTGGTGGCACCCTACGACGGCT  
ACTACCACTGGATCGGGCGGCGTGACCTCGACCAGCAAGACCACCACAACCTGCTAGTAAGACCAG  
CACCCTACGTCCTCGACCTCCTGCACCACCCCCACTGCCGTAGCTGTGACCTTTGATCTGACGG  
CGACCACCACCTACGGCGAGAACATCTACCTGGTCGGGTCGATCTCTCAGCTCGGTGACTGGGAG  
ACCAGCGATGGCATAGCTCTGAGCGCTGACAAGTACACTTCCAGCAACCCGCTTTGGTATGTAAC  
TGTGACTCTGCCGGCTGGTGAGTCATTTGAGTACAAGTTCATCCGCGTCGAGAGCGATGACTCCG  
TGGAGTGGGAGAGCGACCCGAACCGGAATACACCGTTCCTCAGGCGTGCGGCGAGTCGACCGCG  
ACGGTGACCGACACCTGGCGGTAG

FIGURE 6: *Aspergillus awamori* var. *kawachi* GSHE precursor (including the underlined signal sequence and mature protein) protein sequence.

```

M S F R S L L A L S G L V C S G L A S V I S K R A T L D S W L S N
E A T V A R T A I L N N I G A D G A W V S G A D S G I V V A S P S
T D N P D Y F Y T W T R D S G L V I K T L V D L F R N G D T D L L
S T I E H Y I S S Q A I I Q G V S N P S G D L S S G G L G E P K F
N V D E T A Y T G S W G R P Q R D G P A L R A T A M I G F G Q W L
L D N G Y T S A A T E I V W P L V R N D L S Y V A Q Y W N Q T G Y
D L W E E V N G S S F F T I A V Q H R A L V E G S A F A T A V G S
S C S W C D S Q A P Q I L C Y L Q S F W T G S Y I L A N F D S S R
S G K D T N T L L G S I H T F D P E A G C D D S T F Q P C S P R A
L A N H K E V V D S F R S I Y T L N D G L S D S E A V A V G R Y P
E D S Y Y N G N P W F L C T L A A A E Q L Y D A L Y Q W D K Q G S
L E I T D V S L D F F K A L Y S G A A T G T Y S S S S S T Y S S I
V S A V K T F A D G F V S I V E T H A A S N G S L S E Q F D K S D
G D E L S A R D L T W S Y A A L L T A N N R R N S V V P P S W G E
T S A S S V P G T C A A T S A S G T Y S S V T V T S W P S I V A T
G G T T T T A T T T G S G G V T S T S K T T T T A S K T S T T T S
S T S C T T P T A V A V T F D L T A T T T Y G E N I Y L V G S I S
Q L G D W E T S D G I A L S A D K Y T S S N P L W Y V T V T L P A
G E S F E Y K F I R V E S D D S V E W E S D P N R E Y T V P Q A C
G E S T A T V T D T W R

```

Figure 7

